AccessData

Known File Filter (KFF)

Installation Guide

5.6 & 6.0

ACCESSDATA®

AccessData Legal and Contact Information

Document date: December 15, 2015

Legal Information

AccessData Group, Inc. makes no representations or warranties with respect to the contents or use of this documentation, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. Further, AccessData Group, Inc. reserves the right to revise this publication and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes.

Further, AccessData Group, Inc. makes no representations or warranties with respect to any software, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. Further, AccessData Group, Inc. reserves the right to make changes to any and all parts of AccessData software, at any time, without any obligation to notify any person or entity of such changes.

You may not export or re-export this product in violation of any applicable laws or regulations including, without limitation, U.S. export regulations or the laws of the country in which you reside.

AccessData Group, Inc.
1100 Alma Street
Menlo Park, California 94025
USA

AccessData Trademarks and Copyright Information

The following are either registered trademarks or trademarks of AccessData Group, Inc. All other trademarks are the property of their respective owners.

AccessData®	DNA®	PRTK®
AccessData Certified Examiner® (ACE®)	Forensic Toolkit® (FTK®)	Registry Viewer®
AD Summation®	Mobile Phone Examiner Plus®	Summation®
Discovery Cracker®	MPE+ Velocitor™	SilentRunner®
Distributed Network Attack®	Password Recovery Toolkit®	

Contents

Chapter 1

Getting Started with KFF (Known File Filter)

This document contains the following information about understanding and getting started using KFF (Known File Filter).

Important: AccessData applications versions 5.6, 6.0, and later use a new KFF architecture. If you are using one of the following applications version 5.6 or later, you must install and implement the new KFF architecture:

- FTK-based products (FTK, FTK Pro, AD Lab, AD Enterprise)
- Summation
- eDiscovery

See What has Changed in Version 5.6 on page 32.

About KFF

KFF (Known File Filter) is a utility that compares the file hash values of known files against the files in your project. The known files that you compare against may be the following:

- Files that you want to ignore, such as operating system files
- Files that you want to be alerted about, such as malware or other contraband files

The hash values of files, such as MD5, SHA-1, etc., are based on the file's content, not on the file name or extension. The helps you identify files even if they are renamed.

Using KFF during your analysis can provide the following benefits:

- Immediately identify and ignore 40-70% of files irrelevant to the project.
- Immediately identify known contraband files.

Introduction to the KFF Architecture

There are two distinct components of the KFF architecture:

- KFF Data - The KFF data are the hashes of the known files that are compared against the files in your project. The KFF data is organized in KFF Hash Sets and KFF Groups. The KFF data can be comprised of hashes obtained from pre-configured libraries (such as NSRL) or custom hashes that you configure yourself.

  See Components of KFF Data on page 5.
- KFF Server - The KFF Server is the component that is used to store and process the KFF data against your evidence. The KFF Server uses the AccessData Elasticsearch Windows Service. After you install the KFF Server, you import your KFF data into it.

Note: The KFF database is no longer stored in the shared evidence database or on the file system in EDB format.

Components of KFF Data

Item	Description
Hash	The unique MD5 or SHA-1 hash value of a file. This is the value that is compared between known files and the files in your project.
Hash Set	A collection of hashes that are related somehow. The hash set has an ID, status, name, vendor, package, and version. In most cases, a set corresponds to a collection of hashes from a single source that have the same status.
Group	KFF Groups are containers that are used for managing the Hash Sets that are used in a project. KFF Groups can contains Hash Sets as well as other groups. Projects can only use a single KFF Group. However, when configuring your project you can select a single KFF Group which can contains nested groups.
Status	The specified status of a hash set of the known files which can be either Ignore or Alert. When a file in a project matches a known file, this is the reported status of the file in the project.
Library	A pre-defined collection of hashes that you can import into the KFF Serve. There are three pre-defined libraries: ● NSRL ● NDIC HashKeeper ● DHS See About Pre-defined KFF Hash Libraries on page 7.

Item	Description
Index/Indices	When data is stored internally in the KFF Library, it is stored in multiple indexes or indices. The following indices can exist: ● NSRL index A dedicated index for the hashes imported from the NSRL library. ● NDIC index A dedicated index for the hashes imported from the NDIC library. ● DHC index A dedicated index for the hashes imported from the DHC library. ● KFF index A dedicated index for the hashes that you manually create or import from other sources, such as CSV. These indices are internal and you do not see them in the main application. The only place that you see some of them are in the KFF Import Tool. See Using the KFF Import Utility on page 16. The only time you need to be mindful of the indices is when you use the KFF binary format when you either export or import data. See About CSV and Binary Formats on page 22.

About the Organization of Hashes, Hash Sets, and KFF Groups

Hashes, such as MD5, SHA-1, etc., are based on the file's content, not on the file name or extension.

You can also import hashes into the KFF Server in **.CSV** format.

For FTK-based products, you can also import hashes into the KFF Server that are contained in **.TSV**, **.HKE**, **.HKE.TXT**, .HDI, .HDB, **.hash**, **.NSRL**, or .KFF file formats.

You can also manually add hashes.

Hashes are organized into Hash Sets. Hash Sets usually include hashes that have a common status, such as Alert or Ignore.

Hash Sets must be organized into to KFF Groups before they can be utilized in a project.

About Pre-defined KFF Hash Libraries

All of the pre-configured hash sets currently available for KFF come from three federal government agencies and are available in KFF libraries.

See About KFF Pre-Defined Hash Libraries on page 27.

You can use the following KFF libraries:

- NIST NSRL
  See About Importing the NIST NSRL Library on page 19.
- NDIC HashKeeper (Sept 2008)
  See Importing the NDIC Hashkeeper Library on page 20.
- DHS (Jan 2008)
  See Importing the DHS Library on page 21.

It is not required to use a pre-configured KFF library in order to use KFF. You can configure or import custom hash sets. See your application's *Admin Guide* for more information.

How KFF Works

The Known File Filter (KFF) is a body of MD5 and SHA1 hash values computed from electronic files. Some pre-defined data is gathered and cataloged by several US federal government agencies or you can configure you own. KFF is used to locate files residing within project evidence that have been previously encountered by other investigators or archivists. Identifying previously cataloged (known) files within a project can expedite its investigation.

When evidence is processed with the MD5 Hash (and/or SHA-1 Hash) and KFF options, a hash value for each file item within the evidence is computed, and that newly computed hash value is searched for within the KFF data. Every file item whose hash value is found in the KFF is considered to be a known file.

Note: If two hash sets in the same group have the same MD5 hash value, they must have the same metadata. If you change the metadata of one hash set, all hash sets in the group with the same MD5 hash file will be updated to the same metadata.

The KFF data is organized into Groups and stored in the KFF Server. The KFF Server service performs lookup functions.

Status Values

In order to accelerate an investigation, each known file can labeled as either Alert or Ignore, meaning that the file is likely to be forensically interesting (Alert) or uninteresting (Ignore). Other files have a status of Unknown.

The Alert/Ignore designation can assist the investigator to hone in on files that are relevant, and avoid spending inordinate time on files that are not relevant. Known files are presented in the Overview Tab's File Status Container, under "KFF Alert files" and "KFF Ignorable."

Hash Sets

The hash values comprising the KFF are organized into hash sets. Each hash set has a name, a status, and a listing of hash values. Consider two examples. The hash set “ZZ00001 Suspected child porn” has a status of Alert and contains 12 hash values. The hash set “BitDefender Total Security 2008 9843” has a status of Ignore and contains 69 hash values. If, during the course of evidence processing, a file item’s hash value were found to belong to the “ZZ00001 Suspected child porn” set, then that file item would be presented in the KFF Alert files list. Likewise, if another file item’s hash value were found to belong to the “BitDefender Total Security 2008 9843” set, then that file would be presented in the KFF Ignorable list.

In order to determine whether any Alert file is truly relevant to a given project, and whether any Ignore file is truly irrelevant to a project, the investigator must understand the origins of the KFF’s hash sets, and the methods used to determine their Alert and Ignore status assignments.

You can install libraries of pre-defined hash sets or you can import custom hash sets. The pre-defined hash sets contain a body of MD5 and SHA1 hash values computed from electronic files that are gathered and cataloged by several US federal government agencies.

See About KFF Pre-Defined Hash Libraries on page 27.

Higher Level Structure and Usage

Because hash set groups have the properties just described, and because custom hash sets and groups can be defined by the investigator, the KFF mechanism can be leveraged in creative ways. For example, the investigator may define a group of hash sets created from encryption software and another group of hash sets created from child pornography files and then apply only those groups while processing.

About the KFF Server and Geolocation

In order to use the Geolocation Visualization feature in various AccessData products, you must use the KFF architecture and do the following:

- Install the KFF Server.
  See Installing the KFF Server on page 10.
- Install the Geolocation (GeoIP) Data (this data provide location data for evidence)
  See Installing the Geolocation (GeoIP) Data on page 21.
  From time to time, there will be updates available for the GeoIP data.
  See Installing KFF Updates on page 26.

If you are upgrading to 5.6 or later from an application 5.5 or earlier, you must install the new KFF Server and the updated Geolocation data.

Installing the KFF Server

About Installing the KFF Server

In order to use KFF, you must first install and configure an KFF Server.

For product versions 5.6.x and 6.0.x and later, you install a KFF Server by installing the AccessData Elasticsearch Windows Service.

Where you install the KFF Server depends on the product you are using with KFF:

- For FTK and FTK Pro applications, the KFF Server must be installed on the same computer that runs the FTK Examiner application.
- For all other applications, such as AD Lab, Summation, or eDiscovery, the KFF Server can be installed on either the same computer as the application or on a remote computer. For large environments, it is recommended that the KFF Server be installed on a dedicated computer.

After installing the KFF Server, you configure the application with the location of the KFF Server.

See Configuring the Location of the KFF Server on page 12.

About KFF Server Versions

The KFF Server (AccessData Elasticsearch Windows Service) may be updated from time to time. It is best to use the latest version.

AccessData Elasticsearch Windows Service	Released	Installation Instructions

Version 1.3.2.x	• November 2014 with 5.6 versions of
	  ▪ FTK-based products
	  ▪ Summation
	  ▪ eDiscovery
	• November 2015 with 6.0 versions of
	  ▪ FTK-based products
	  ▪ Summation
	  ▪ eDiscovery
		See Installing the KFF Server Service on page 11.

For applications 5.5 and earlier, the KFF Server component was version 1.2.7 and earlier.

About Upgrading from Earlier Versions

If you have used KFF with applications versions 5.5 and earlier, you can migrate your legacy KFF data to the new architecture.

See Migrating Legacy KFF Data on page 13.

Process for Installing KFF

The process for installing KFF is as follows:

1. Downloading the Latest KFF Installation Files (page 11)
2. Installing the KFF Server Service (page 11)
3. Configuring the KFF Server location:
   - Configuring the KFF Server Location on FTK-based Computers (page 12)
   - Configuring the KFF Server Location on Summation and eDiscovery Applications (page 12)
4. (Optional) Upgrading or importing KFF data.
   - See Migrating Legacy KFF Data on page 13.
   - About Importing KFF Data (page 15)
   - Importing Pre-defined KFF Data Libraries (page 18)
   - Installing the Geolocation (GeoIP) Data (page 21)

Downloading the Latest KFF Installation Files

You can download ISO files which has the latest KFF files. Files may be updated from time to time.

To download the latest KFF Installation Files

1. Go to the AccessData Current Releases - Digital Forensics product download page.
   You can also download the file from the FTK or AD Lab product download pages.
2. Click **Known File Filter (KFF) Compatible with 5.6 and above**.
3. Do one of the following:
   - To download the KFF Server files, utilities, and NSRL data, click **KFF for all 6.0 products**.
   - To download the DHS library, click **KFF DHS**.
   - To download the NDIC library, click **KFF NDIC**.
4. Click **Download Now.**

Installing the KFF Server Service

The KFF Server Service is install by installing the AccessData Elasticsearch Windows Service

For instructions on installing the AccessData Elasticsearch Windows Service, see Installing the Elasticsearch Service (page 34).

Configuring the Location of the KFF Server

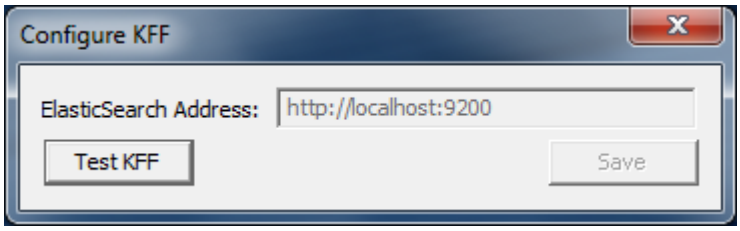
After installing the KFF Server, on the computer running the application, such as FTK, AD Lab, Summation, or eDiscovery, you configure the location of the KFF Server.

Do one of the following:

- Configuring the KFF Server Location on FTK-based Computers (page 12)
- Configuring the KFF Server Location on Summation and eDiscovery Applications (page 12)

Configuring the KFF Server Location on FTK-based Computers

Before using KFF with FTK, FTK Pro, Lab, or Enterprise, with KFF, you must configure the location of the KFF Server.



Important: To configure KFF, you must be logged in with Admin privileges.

To view or edit KFF configuration settings

1. In the *Case Manager*, click **Tools > Preferences > Configure KFF**.
2. You can set or view the address of the KFF Server.
   - If you installed the KFF Server on the same computer as the application, this value will be localhost.
   - If you installed the KFF Server on a different computer, identify the KFF server.
3. Click **Test** to validate communication with the KFF Server.
4. Click **Save**.
5. Click **OK**.

Configuring the KFF Server Location on Summation and eDiscovery Applications

When using the KFF Server with Summation or eDiscovery applications, two configuration files must point to the KFF Server location.

These setting are configured automatically during the KFF Server installation. If needed, you can verify the settings.

However, if you change the location of the KFF Server, do the following to specify the location of the KFF Server.

1. Configure **AdgWindowsServiceHost.exe.config:**
   1a. On the computer running the application (for example, the server running Summation), go to **C:\Program Files\AccessData\Common\FTK Business Services**.
   1b. Open **AdgWindowsServiceHost.exe.config**.

1c. Modify the line **<add key="KffElasticSearchUrl" value="http://localhost:9200" />**.

1d. Change *localhost* to be the location of your KFF server (you can use hostname or IP).

1e. Save and close file.

1f. Restart the business services common service.

2. Configure AsyncProcessingServices **web.config**:

2a. On the computer running the application (for example, the server running Summation), go to **C:\Program Files\AccessData\AsyncProcessingServices**.

2b. Open **web.config**.

2c. Modify the line **<add key="KffElasticSearchUrl" value="http://localhost:9200" />**.

2d. Change *localhost* to be the location of your KFF server (you can use hostname or IP).

2e. Save and close file.

2f. Restart the AsyncProcessing service.

Migrating Legacy KFF Data

If you have used KFF with applications versions 5.5 and earlier, you can migrate that data from the legacy KFF Server to the new KFF Server architecture.

Important: Applications version 5.6 and later can only use the new KFF architecture that was introduced in 5.6. If you want to use KFF data from previous versions, you must migrate the data.

Important: If you have NSRL, NDIC, or DHS data in your legacy data, those sets will not be migrated. You must re-import them using the 5.6 versions or later of those libraries. Only legacy custom KFF data will be migrated.

Legacy KFF data is migrated to KFF Groups and Hash Sets on the new KFF Server.

Because KFF Templates are no longer used, they will be migrated as KFF Groups, and the groups that were under the template will be added as sub-groups.

You migrate data using the KFF Migration Tool. To use the KFF Migration Tool, you identify the following:

- The Storage Directory folder where the legacy KFF data is located.

  This was folder was configured using the KFF Server Configuration utility when you installed the legacy KFF Server. If needed, you can use this utility to view the KFF Storage Directory. The default location of the KFF_Config.exe file is Program Files\AccessData\KFF.

- The URL of the new KFF Server (the computer running the AccessData Elastic Search Windows Service)

  This is populated automatically if the new KFF Server has been installed.

To install the KFF Migration Tool

1. On the computer where you have installed the KFF Server, access the KFF Installation disc, and run the **autorun.exe**.
2. Click the *64 bit* or *32 bit* **Install KFF Migration Utility**.
3. Complete the installation wizard.

To migrate legacy KFF data

1. On the legacy KFF Server, you must stop the KFF Service.

   You can stop the service manually or use the legacy KFF Config.exe utility.

2. On the new KFF Server, launch the KFF Migration Tool.
3. Enter the directory of the legacy KFF data.
4. The URL of Elasticsearch should be listed.
5. Click **Start**.
6. When completed, review the summary data.

Importing KFF Data

About Importing KFF Data

You can import hashes and KFF Groups that have been previous configured.

You can import KFF data in one of the following formats:

KFF Data sources that you can import

	Description
Pre-configured KFF libraries	You can import KFF data from the following pre-configured libraries ● NIST NSRL ● NDIC HashKeeper ● DHS To import KFF libraries, it is recommended that you use the KFF Import Utility. See Using the KFF Import Utility on page 16. See Importing Pre-defined KFF Data Libraries on page 18. See KFF Library Reference Information on page 27.
Custom Hash Sets and KFF Groups	You can import custom hashes from CSV files. See About the CSV Format on page 22. For FTK-based products, you can also import custom hashes from the following file types: ● Delimited files (CSV or TSV) ● Hash Database files (HDB) ● Hashkeeper files (HKE) ● FTK Exported KFF files (KFF) ● FTK Supported XML files (XML) ● FTK Exported Hash files (HASH) To import these kinds of files, use the KFF Import feature in your application. See *Using the Known File Feature* chapter.
KFF binary files	You can import KFF data that was exported in a KFF binary format, such as an archive of a KFF Server. See About CSV and Binary Formats on page 22. When you import a KFF binary snapshot, you must be running the same version of the KFF Server as was used to create the binary export. To import KFF binary files, it is recommend that you use the KFF Import Utility. See Using the KFF Import Utility on page 16.

About KFF Data Import Tools

When you import KFF data, you can use one of two tools:

KFF Data Import Tools

The application's Import feature	The KFF management feature in the application lets you import both .CSV and KFF Binary formats. Use the application to import .CSV files. See *Using the Known File Feature* chapter. Even though you can import KFF binary files using the application, it is recommend that you use the KFF Import Utility.
KFF Import Utility	It is recommended that you use the KFF Import Utility to import KFF binary files. See Using the KFF Import Utility on page 16.

About Default Status Values

When you import KFF data, you configure a default status value of Alert or Ignore. When adding Hash Sets to KFF Groups, you can configure the KFF Groups to use the default status values of the Hash Set or you can configure the KFF Group with a status that will override the default Hash Set values.

See Components of KFF Data on page 5.

About Duplicate Hashes

If multiple Hash Set files containing the same Hash identifier are imported into a single KFF Group, the group keeps the last Hash Set's metadata information, overwriting the previous Hash Sets' metadata. This only happens within an individual group and not across multiple groups.

Using the KFF Import Utility

About the KFF Import Utility

Due to the large size of some KFF data, a stand-alone KFF Import utility is available to use to import the data. This KFF Import utility can import large amounts of data faster then using the import feature in the application.

It is recommend that you install and use the KFF Import utility to import the following:

- NSRL, DHC, and NIST libraries
- An archive of a KFF Server that was exported in the binary format

After importing NSRL, NDIC, or DHS libraries, these indexes are displayed in the *Currently Installed Sets* list.

See Components of KFF Data on page 5.

You can also use the KFF Import Utility to remove the NSRL, NDIC, or DHS indexes that you have imported.

An archive of a KFF Server, which is the exported *KFF Index*, is not shown in the list.

Installing the KFF Import Utility

You should use the KFF Import Utility to import some kinds of KFF data.

To install the KFF Import Utility

1. On the computer where you have installed the KFF Server, access the KFF Installation disc, and run the `autorun.exe`.
2. Click the *64 bit* or *32 bit* **Install KFF Import Utility**.
3. Complete the installation wizard.

Importing a KFF Server Archive Using the KFF Import Utility

You can import an archive of a KFF Server that you have exported using the binary format.

If you are importing a pre-defined KFF Library, see Importing Pre-defined KFF Data Libraries (page 18).

To import using the KFF Import Utility

1. On the KFF Server, open the KFF Import Utility.
2. To test the connection to the KFF Server's Elasticsearch service at the displayed URL, click **Connect**.
   If it connects correctly, no error is shown.
   If it is not able to connect, you will get the following error: Failed after retrying 10 times: 'HEAD accessdata_threat_indicies'.
3. To import, click **Import**.
4. Click **Browse**.
5. Browse to the folder that contains the KFF binary files.
   Specifically, select the folder that contains the Export.xml file.
6. Click **Start**.
7. Close the dialog.

Removing Pre-defined KFF Libraries Using the KFF Import Utility

You can remove a pre-defined KFF Library that you have previously imported.

You cannot see or remove existing custom KFF data (the *KFF Index*).

To remove pre-defined KFF Libraries

1. On the KFF Server, open the KFF Import Utility.
2. Select the library that you want to remove.
3. Click **Remove**.

Importing Pre-defined KFF Data Libraries

About Importing Pre-defined KFF Data Libraries

After you install the KFF Server, you can import pre-defined NIST NSRL, NDIC HashKeeper, and DHS data libraries.

See About Pre-defined KFF Hash Libraries on page 7.

In versions 5.5 and earlier, you installed these using an executable file. In versions 5.6 and later, you must import them. It is recommend that you use the KFF Import Utility.

After importing pre-defined KFF Libraries, you can remove them from the KFF Server.

See Removing Pre-defined KFF Libraries Using the KFF Import Utility on page 17.

See the following sections:

- About Importing the NIST NSRL Library (page 19)
- Importing the NDIC Hashkeeper Library (page 20)
- Importing the DHS Library (page 21)

About Importing the NIST NSRL Library

You can import the NSRL library into your KFF Server. During the import, two KFF Groups are created: NSRL_Alert and NSRL_Ignore. In FTK-based products, these two groups are automatically added to the Default KFF Group.

The NSRL libraries are updated from time to time. To import and maintain the NSRL data, you do the following:

Process for Importing and Maintaining the NIST NSRL Library

1. Import the complete NSRL library.	You must first install the most current complete NSRL library. You can later add updates to it. To access and import the complete NSRL library, see Importing the Complete NSRL Library (page 20)
2. Import updates to the library	When updates are made available, import the updates to bring the data up-to date. See Installing KFF Updates on page 26. **Important:** In order to use the NSRL updates, you must first import the complete library. When you install an NSRL update, you must keep the previous NSRL versions installed in order to maintain the complete set of NSRL data.

Available NRSL library files (new format)

NSRL Library Release	Released	Information
Complete library version 2.45 (source .ZIP file)	Nov 2014	For use only with applications version 5.6 and later. Contains the full NSRL library up through update 2.45. See Importing the Complete NSRL Library on page 20.

Available Legacy NRSL library files

Legacy NSRL Library Release	Released	Information
version 2.44 (.EXE file)	Nov 2013	For use with the legacy KFF Server that was used with applications versions 5.5 and earlier. Contains the full NSRL library up through update 2.44. Install this library first. **Note:** NSRL updates for the legacy KFF format will end in the 2nd quarter of 2015. From that time, NSRL updates will only be provided in the new format.

Importing the Complete NSRL Library

To add the NSRL library to your KFF Library, you import the data. You start by importing the full NSRL library. You can then import any updates as they are available.

See About Importing the NIST NSRL Library on page 19.

See Installing KFF Updates on page 26.

Important: The complete NSRL library data is contained in a large (3.4 GB) .ZIP file. When expanded, the data is about 18 GB. Make sure that your file system can support files of this size.

Important: Due to the large amount of NSRL data, it will take 3-4 hours to import the NSRL data using the KFF Import Utility. If you import from within an application, it will take even longer.

To install the NSRL complete library

1. Extract the NSRLSOURCE_2.45.ZIP file from the KFF Installation disc.
   See Downloading the Latest KFF Installation Files on page 11.
2. On the KFF Server, launch the *KFF Import Utility*.
   See Installing the KFF Import Utility on page 17.
3. Click **Import**.
4. Click **Browse**.
5. Browse to and select the NSRLSource_2.45 folder that contains the **NSRLFile.txt** file.
   (Make sure you are selecting the folder and not drilling into the folder to select an individual file. The import process will drill into the folder to get the proper files for you.)
6. Click **Select Folder**.
7. Click **Start**.
8. When the import is complete, click **OK**.
9. Close the *Import Utility* dialog and the NSRL library will be listed in the *Currently Installed Sets*.

Importing the NDIC Hashkeeper Library

You can import the Hashkeeper 9.08 library.

For application versions 5.6 and later, these files are stored in the KFF binary format.

To import the Hashkeeper library

1. Have access the NDIC source files by download the ZIP file from the web:
   See Downloading the Latest KFF Installation Files on page 11.
2. Extract the ZIP file.
3. On the KFF Server, launch the *KFF Import Utility*.
   See Installing the KFF Import Utility on page 17.
4. Click **Import**.
5. Click **Browse**.
6. Browse to and select the NDIC source folder that contains the **Export.xml** file.
   (Make sure you are selecting the folder and not drilling into the folder to select an individual file. The import process will drill into the folder to get the proper files for you.)
7. Click **Select Folder**.

8. Click **Start**.
9. When the import is complete, click **OK**.
10. Close the *Import Utility* dialog and the NDIC library will be listed in the *Currently Installed Sets*.

Importing the DHS Library

You can import the DHS 1.08 library.

For application versions 5.6 and later, these files are stored in the KFF binary format.

To import the DHS library

1. Have access the NDIC source files by download the ZIP file from the web:
   See Downloading the Latest KFF Installation Files on page 11.
2. Extract the ZIP file.
3. On the KFF Server, launch the *KFF Import Utility*.
   See Installing the KFF Import Utility on page 17.
4. Click **Import**.
5. Click **Browse**.
6. Browse to and select the DHS source folder that contains the **Export.xml** file.
   (Make sure you are selecting the folder and not drilling into the folder to select an individual file. The import process will drill into the folder to get the proper files for you.)
7. Click **Select Folder**.
8. Click **Start**.
9. When the import is complete, click **OK**.
10. Close the *Import Utility* dialog and the DHS library will be listed in the *Currently Installed Sets*.

Installing the Geolocation (GeoIP) Data

Geolocation (GeoIP) data is used for the Geolocation Visualization feature of several AccessData products.

See About the KFF Server and Geolocation on page 9.

You can also check for and install GeoIP data updates.

If you are upgrading to 5.6 or later from an application 5.5 or earlier, you must install the new KFF Server and the updated Geolocation data.

The Geolocation data that was used with versions 5.5 and earlier is version 1.0.1 or earlier.

The Geolocation data that is used with versions 5.6 and later is version 2014.10 or later.

To install the Geolocation IP Data

1. On the computer where you have installed the KFF Server, access the KFF Installation disc, and run the **autorun.exe**.
   See Downloading the Latest KFF Installation Files on page 11.
2. Click the *64 bit* or *32 bit* **Install Geolocation Data**.
3. Complete the installation wizard.

About CSV and Binary Formats

When you export and import KFF data, you can use one of two formats:

- CSV
- KFF Binary

About the CSV Format

When you use the .CSV format, you use a single .CSV file. The .CSV file contains the hashes that you import or export.

When you export to a CSV file, it contains the hashes as well as all of the information about any associated Hash Sets and KFF Groups. You can only use the CSV format when exporting individual Hash Sets and KFF Groups.

When you import using a CSV file, it can be a simple file containing only the hashes of files, or it can contain additional information about Hash Sets and KFF Groups.

However, CSV files will usually take a little longer to export and import.

To view the sample of a .CSV file that contains binaries and Hash Sets and KFF Groups, perform a CSV export and view the file in Excel.

You can also use the format of CSV files that were exported in previous versions.

To import .CSV files, use the application's KFF Import feature.

About the KFF Binary Format

When you use the KFF binary format, you use a set of files that are in an internal KFF Server (Elasticsearch) format that is referred to as a Snapshot. The binary format is essentially a snapshot of one of the indices contained in the KFF Server. You can only have one binary format snapshot for each index.

See Components of KFF Data on page 5.

The benefit of the binary format is that it is able to support larger amounts of data than the CSV format. For large data sets, the binary format will export and import faster than the CSV format.

For example, when you import the DHC or NDIC Hashkeeper libraries, they are imported from a KFF binary format.

If you export your custom Hash Sets or KFF Groups using the KFF binary format, everything in the *KFF Index* is included.

See About Choosing to Export in CSV or KFF Binary Format on page 23.

When exporting in a Binary format, you specify an existing parent folder and then the name of a new sub-folder for the binary data. The new sub-folder must not previously exist and will be created by the export process.

After export, the binary export folder contains the following:

- **Indices** sub-folder - The folder contains the exported KFF data
- **Export.xml** - This file is the only file that is not an Elasticsearch file and is created by the export feature and contains the KFF Group and Hash Set definitions for the index.

- **Index** - an index file generated by Elasticsearch
- **metadata-snaphot** file with the data and time it was created
- **snapshot-snaphot** file with the data and time it was created

Note: The binary format is dependent on the version of the KFF Server. When exporting and importing the binary format, the systems must be using the same version of the KFF Server.
When new versions of the KFF Server are released in the future, an upgrade process will also be provided.

About Choosing to Export in CSV or KFF Binary Format

When you export your own KFF data, you have the option of using either the CSV or the binary format. The results are different based on the format that you use:

CSV format	
Exporting in CSV format	When you export KFF data using the CSV format, you can export specific pieces of KFF data, such as one or more Hash Sets or one or more KFF Groups. The exported data is contained in one .CSV file. The benefits of the CSV format are that CSV files can be easily viewed and can be manually edited. They are also less dependent on the version of the KFF Server.
Importing from CSV format	When you import a CSV file, the data in the file is data is added to your existing KFF data that is in the *KFF Index*. See Components of KFF Data on page 5. For example, suppose you started by manually created four Hash Sets and one KFF Group. That would be the only contents in your *KFF Index*. Suppose you import a .CSV file that contains five hash sets and two KFF Groups. They will be added together for a total of nine Hash Sets and three KFF Groups. To import .CSV files, use the KFF Import feature in your application. See *Using the Known File Feature* chapter.
KFF binary format	
Exporting in KFF binary format	If you export your KFF data using the KFF binary format, all of the data that you have in the *KFF Index* will be exported together. You cannot use this format to export individual Hash Sets or KFF Groups. See Components of KFF Data on page 5. You will only want to use this format if you intend to export all of the data in the *KFF Index* and import it as a whole. This can be useful in making an archive of your KFF data or copying KFF data from one KFF Server to another. Because NSRL, NIST, and DHC data is contained in their own indexes, when you do an export using this format, those sets are not included. Only the data in the *KFF Index* is exported.

Importing KFF binary format	**IMPORTANT:** When you import a KFF binary format, it will import the complete index and will *replace* any data that is currently in that index on the KFF Server. For example, if you import the DHC library, and then later you import the DHC library again, the DHC index will be replaced with the new import. If you have a KFF binary format snapshot of custom KFF data (which would have come from a binary format export) it will replace all KFF data that already exists in your *KFF Index*. For example, suppose you manually created four Hash Sets and one KFF Group. Suppose you then import a binary format that has five hash sets and two KFF Groups. The binary format will be imported as a complete index and will replace the existing data. The result will be only be the imported five Hash Sets and two KFF libraries. When importing KFF binary files, it is recommend that you use the KFF Import Utility. See Installing the KFF Import Utility on page 17.

Uninstalling KFF

You can uninstall KFF application components independently of the KFF Data.

Main version	Description
Applications 5.6 and later	For applications version 5.6 and later, you uninstall the following components: • *AccessData Elasticsearch Windows Service* (KFF Server) v1.2.7 and later Note: Elasticsearch is used by multiple features in various applications, use caution when uninstalling this service or the related data. • *AccessData KFF Import Utility* (v5.6 and later) • *AccessData KFF Migration Tool* (v1.0 and later) • *AccessData Geo Location Data* (v2014.10 and later) Note: This component is not used by the KFF feature, but with the KFF Server for the geolocation visualization feature. The location of the KFF data is configured when the *AccessData Elasticsearch Windows Service* was installed. By default, it is lactated at C:\Program Files\AccessData\Elacticsearch\Data.
Applications 5.5 and earlier	• KFF Server (v1.2.7 and earlier) Note: The KFF Server is also used by the geolocation visualization feature. • AccessData Geo Location Data (1.0.1 and earlier) This component is not used by the KFF feature, but with the KFF Server for the geolocation visualization feature. The location of the KFF data was configured when the *KFF Server* was installed. You can view the location of the data by running the *KFF.Config.exe* on the KFF Server. If you are upgrading from 5.5 to 5.6, you can migrate your KFF data before uninstalling the KFF Server.

Installing KFF Updates

From time to time, AccessData will release updates to the KFF Server and the KFF data libraries.

Some of the KFF data updates may require you to update the version of the KFF Server.

To check for updates, do the following:

1. Go to the KFF product download page.
   See Downloading the Latest KFF Installation Files on page 11.
2. Check for updates.
   - See About KFF Server Versions on page 10.
   - See About Importing the NIST NSRL Library on page 19.
3. If there are updates, download them.
4. Install or import the updates.

KFF Library Reference Information

About KFF Pre-Defined Hash Libraries

This section includes a description of pre-defined hash collections that can be added as AccessData KFF data.

The following pre-defined libraries are currently available for KFF and come from one of three federal government agencies:

- NIST NSRL (The default library installed with KFF)
- NDIC HashKeeper (An optional library that can be downloaded from the AccessData Downloads page)
- DHS (An optional library that can be downloaded from the AccessData Downloads page)

Note: Because KFF is now multi-sourced, it is no longer maintained in HashKeeper format. Therefore, you cannot modify KFF data in the HashKeeper program. However, the HashKeeper format continues to be compatible with the AccessData KFF data.

Use the following information to help identify the origin of any hash set within the KFF

- The NSRL hash sets do not begin with “ZZN” or “ZN”. In addition, in the AD Lab KFF, all the NSRL hash set names are appended (post-fixed) with multi-digit numeric identifier. For example: “Password Manager & Form Filler 9722.”
- All HashKeeper Alert sets begin with “ZZ”, and all HashKeeper Ignore sets begin with “Z”. (There are a few exceptions. See below.) These prefixes are often followed by numeric characters (“ZZN” or “ZN” where N is any single digit, or group of digits, 0-9), and then the rest of the hash set name. Two examples of HashKeeper Alert sets are:
  - “ZZ00001 Suspected child porn”
  - “ZZ14W”

  An example of a HashKeeper Ignore set is:
  - “Z00048 Corel Draw 6”
- The DHS collection is broken down as follows:
  - In 1.81.4 and later there are two sets named “DHS-ICE Child Exploitation JAN-1-08 CSV” and “DHS-ICE Child Exploitation JAN-1-08 HASH”.
  - In AD Lab there is just one such set, and it is named “DHS-ICE Child Exploitation JAN-1-08”.

Once an investigator has identified the vendor from which a hash set has come, he/she may need to consider the vendor’s philosophy on collecting and categorizing hash sets, and the methods used by the vendor to gather hash values into sets, in order to determine the relevance of Alert (and Ignore) hits to his/her project. The following descriptions may be useful in assessing hits.

NIST NSRL

The NIST NSRL collection is described at: http://www.nsrl.nist.gov/index.html. This collection is much larger than HashKeeper in terms of the number of sets and the total number of hashes. It is composed entirely of hash sets being generated from application software. So, all of its hash sets are given Ignore status by AccessData staff except for those whose names make them sound as though they could be used for illicit purposes.

The NSRL collection divides itself into many sub-collections of hash sets with similar names. In addition, many of these hash sets are “empty”, that is, they are not accompanied by any hash values. The size of the NSRL collection, combined with the similarity in set naming and the problem of empty sets, allows AccessData to modify (or selectively alter) NSRL’s own set names to remove ambiguity and redundancy.

Find contact info at http://www.nsrl.nist.gov/Contacts.htm.

NDIC HashKeeper

NDIC’s HashKeeper collection uses the Alert/Ignore designation. The Alert sets are hash values contributed by law enforcement agents working in various jurisdictions within the US - and a few that apparently come from Luxemburg. All of the Alert sets were contributed because they were believed by the contributor to be connected to child pornography. The Ignore sets within HashKeeper are computed from files belonging to application software.

During the creation of KFF, AccessData staff retains the Alert and Ignore designations given by the NDIC, with the following exceptions. AccessData labels the following sets Alert even though HashKeeper had assigned them as Ignore: “Z00045 PGP files”, “Z00046 Steganos”, “Z00065 Cyber Lock”, “Z00136 PGP Shareware”, “Z00186 Misc Steganography Programs”, “Z00188 Wiping Programs”. The names of these sets may suggest the intent to conceal data on the part of the suspect, and AccessData marks them Alert with the assumption that investigators would want to be “alerted” to the presence of data obfuscation or elimination software that had been installed by the suspect.

The following table lists actual HashKeeper Alert Set origins:

A Sample of HashKeeper KFF Contributions

Hash	Contributor	Location	Contact Information	Case/Source
ZZ00001 Suspected child porn	Det. Mike McNown & Randy Stone	Wichita PD		
ZZ00002 Identified Child Porn	Det. Banks	Union County (NJ) Prosecutor's Office	(908) 527-4508	case 2000S-0102
ZZ00003 Suspected child porn	Illinois State Police			
ZZ00004 Identified Child Porn	SA Brad Kropp, AFOSI, Det 307		(609) 754-3354	Case # 00307D7-S934831

A Sample of HashKeeper KFF Contributions (Continued)

Hash	Contributor	Location	Contact Information	Case/Source
ZZ00000, suspected child porn	NDIC			
ZZ00005 Suspected Child Porn	Rene Moes, Luxembourg Police		rene.moes@police.etat.lu	
ZZ00006 Suspected Child Porn	Illinois State Police			
ZZ00007b Suspected KP (US Federal)				
ZZ00007a Suspected KP Movies				
ZZ00007c Suspected KP (Alabama 13A-12-192)				
ZZ00008 Suspected Child Pornography or Erotica	Sergeant Purcell	Seminole County Sheriff's Office (Orlando, FL, USA)	(407) 665-6948, dpurcell@seminolesheriff.org	suspected child pornogrpahy from 20010000850
ZZ00009 Known Child Pornography	Sergeant Purcell	Seminole County Sheriff's Office (Orlando, FL, USA)	(407) 665-6948, dpurcell@seminolesheriff.org	200100004750
ZZ10 Known Child Porn	Detective Richard Voce CFCE	Tacoma Police Department	(253)594-7906, rvoce@ci.tacoma.wa.us	
ZZ00011 Identified CP images	Detective Michael Forsyth	Baltimore County Police Department	(410)887-1866, mick410@hotmail.com	
ZZ00012 Suspected CP images	Sergeant Purcell	Seminole County Sheriff's Office (Orlando, FL, USA)	(407) 665-6948, dpurcell@seminolesheriff.org	
ZZ0013 Identified CP images	Det. J. Hohl	Yuma Police Department	928-373-4694	YPD02-70707

A Sample of HashKeeper KFF Contributions (Continued)

Hash	Contributor	Location	Contact Information	Case/Source
ZZ14W	Sgt Stephen May		Tamara.Chandler@oag.state.tx.us, (512)936-2898	TXOAG 41929134
ZZ14U	Sgt Chris Walling		Tamara.Chandler@oag.state.tx.us, (512)936-2898	TXOAG 41919887
ZZ14X	Sgt Jeff Eckert		Tamara.Chandler@oag.state.tx.us, (512)936-2898	TXOAG Internal
ZZ14I	Sgt Stephen May		Tamara.Chandler@oag.state.tx.us, (512)936-2898	TXOAG 041908476
ZZ14B	Robert Britt, SA, FBI		Tamara.Chandler@oag.state.tx.us, (512)936-2898	TXOAG 031870678
ZZ14S	Sgt Stephen May		Tamara.Chandler@oag.state.tx.us, (512)936-2898	TXOAG 041962689
ZZ14Q	Sgt Cody Smirl		Tamara.Chandler@oag.state.tx.us, (512)936-2898	TXOAG 041952839
ZZ14V	Sgt Karen McKay		Tamara.Chandler@oag.state.tx.us, (512)936-2898	TXOAG 41924143
ZZ00015 Known CP Images	Det. J. Hohl	Yuma Police Department	928-373-4694	YPD04-38144
ZZ00016	Marion County Sheriff's Department		(317) 231-8506	MP04-0216808

The basic rule is to always consider the source when using KFF in your investigations. You should consider the origin of the hash set to which the hit belongs. In addition, you should consider the underlying nature of hash values in order to evaluate a hit's authenticity.

Higher Level KFF Structure and Usage

Since hash set groups have the properties just described (and because custom hash sets and groups can be defined by the investigator) the KFF mechanism can be leveraged in creative ways. For example:

- You could define a group of hash sets created from encryption software and another group of hash sets created from child pornography files. Then, you would apply only those groups while processing.
- You could also use the Ignore status. You are about to process a hard drive image, but your search warrant does not allow inspection of certain files within the image that have been previously identified. You could do the following and still observe the warrant:
  - 4a. Open the image in Imager, navigate to each of the prohibited files, and cause an MD5 hash value to be computed for each.
  - 4b. Import these hash values into custom hash sets (one or more), add those sets to a custom group, and give the group Ignore status.
  - 4c. Process the image with the MD5 and KFF options, and with AD_Alert, AD_Ignore, and the new, custom group selected.
  - 4d. During post-processing analysis, filter file lists to eliminate rows representing files with Ignore status.

Hash Set Categories

The highest level of the KFF's logical structure is the categorizing of hash sets by owner and scope. The categories are AccessData, Project Specific, and Shared.

Hash Set Categories

	Description
AccessData	The sets shipped with as the Library. Custom groups can be created from these sets, but the sets and their status values are read only.
Project Specific	Sets and groups created by the investigator to be applied only within an individual project.
Shared	Sets and groups created by the investigator for use within multiple projects all stored in the same database, and within the same application schema.

Important: Coordination among other investigators is essential when altering Shared groups in a lab deployment. Each investigator must consider how other investigators will be affected when Shared groups are modified.

What has Changed in Version 5.6

WIth the 5.6 release of eDiscovery, Summation, and FTK-based products, the KFF feature has been updated.

If you used KFF with applications version 5.5 or earlier, you will want to be aware of the following changes in the KFF functionality.

Changes from version 5.5 to 5.6

Item	Description
KFF Server	KFF Server now runs a different service. • In 5.5 and earlier, the KFF Server ran as the *KFF Server* service. • In 5.6 and later, the KFF Server uses the *AccessData Elasticsearch Windows Service*. For applications version 5.6 and later, all KFF data must be created in or imported into the new KFF Server.
KFF Migration Tool	This is a new tool that lets you migrate custom KFF data from 5.5 and earlier to the new KFF Server. NIST NSRL, NDIC HashKeeper, or DHS library data from 5.5 will not be migrated. You must re-import it. See Migrating Legacy KFF Data on page 13.
KFF Import Utility	This is a new utility that lets you import large amounts of KFF data quicker than using the import feature in the application. See Using the KFF Import Utility on page 16.
KFF Libraries, Templates, and Groups	In 5.5, all Hash Sets were configured within KFF Libraries. KFF Libraries could then contain KFF Groups and KFF Templates. KFF Libraries and Templates have been eliminated. You now simply create or import KFF Groups and add Hash Sets to the groups. You can now nest KFF Groups.
NIST NSRL, NDIC HashKeeper, or DHS libraries	In 5.5 and earlier, to use these libraries, you ran an installation wizard for each library. You now import these libraries using the KFF Import Utility. See About Importing Pre-defined KFF Data Libraries on page 18.
Import Log	FTK-based products no longer include the Import Log. eDiscovery and Summation products did not have it previously.
Export	When you export KFF data you can now choose two formats: • CSV format which replaced XML format • A new binary format See About CSV and Binary Formats on page 22.

Chapter 2

Installing the AccessData Elasticsearch Windows Service

About the Elasticsearch Service

The AccessData Elasticsearch Windows Service is used by multiple features in multiple applications, including the following:

- KFF (Known File Filter) in all applications
- Visualization Geolocation in all applications

The AccessData Elasticsearch Windows Service uses the Elasticsearch open source search engine.

Prerequisites

- For best results with eDiscovery products and AD Lab and Enterprise, you should install the AccessData Elasticsearch Windows Service on a dedicated computer that is different from the computer running the application that uses it.

  For single-computer installations such as FTK, you can install the AccessData Elasticsearch Windows Service on the same computer as the application.

  A single instance of an AccessData Elasticsearch Windows Service is usually sufficient to support multiple features. However, if your network is extensive, you may want to install the service on multiple computers on the network. Consult with support for the best configuration for your organization's network.
- You can install the AccessData Elasticsearch Windows Service on 32-bit or 64-bit computers.
- 16 GB of RAM or higher
- Microsoft .NET Framework 4

  To install the AccessData Elasticsearch Windows Service, Microsoft .NET Framework 4 is required. If you do not have .NET installed, it will be installed automatically.
- If you install the AccessData Elasticsearch Windows Service on a system that has not previously had an AccessData product installed upon it, you must add a registry key to the system in order for the service to install correctly.

Installing the Elasticsearch Service

Installing the Service

To install the AccessData Elasticsearch Windows Service

1. Click the the AccessData Elasticsearch Windows Service installer.

   It is avaialable on the KFF Installation disc by clicking *autorun.exe*.

2. Accept the License Agreement and click **Next**.
3. On the *Destination Folder* dialog, click **Next** to install to the folder, or click **Change** to install to a different folder.

   This is where the Elasticsearch folder with the Elasticsearch service is installed.

4. On the *Data Folder* dialog, click **Next** to install to the folder, or click **Change** to install to a different folder.

   This is where the Elasticsearch data is stored.

   Note: This folder may contain up to 10GB of data.

5. (For use with KFF) In the *User Credentials* dialog, you can configure credentials to access KFF Data files that you want to import if they exist on a different computer.

   This provides the credentials for the Elasticsearch service to use in order to access a network share with a user account that has permissions to the share.

   Enter the user name, the domain name, and the password. If the user account is local, do not enter any domain value, such as localhost. Leave it blank instead.

6. In the *Allow Remote Communication* dialog, enter the IP address(es) of any machine(s) that will have ThreatBridge installed. If you plan on installing ThreatBridge on the same server as the AccessData Elasticsearch Windows Service, click **Next**.
7. *Select Enable Remote Communication*.

   Note: If Enable Remote Communication is selected, a firewall rule will be created to allow communication to the AccessData Elasticsearch Windows Service service for every IP address added to the IP Address field. If no IP addresses are listed, then ANY IP address will be able to access the AccessData Elasticsearch Windows Service.

8. In the following *Allow Remote Communication* dialog, accept the default HTTP and Transport TCP Port values and click **Next**. However, if there are conflicts with these ports on the network, change the values to use other ports.
9. The *Configuration 1* dialog contains the following fields:
   - **Cluster name** - This field automatically populates with the system's name.
   - **Node name** - This field automatically populates with the system's name.

     Note: If installing the AccessData Elasticsearch Windows Service on more than one system, allow the first system to install with the system's name in the cluster and the node fields. In the second and subsequent systems, enter the first system's name in the cluster field, and in the node field, enter the name of the system to which you are installing.

   - **Heap size** - This is the memory allocated for the AccessData Elasticsearch Windows Service. Normally you can accept the default value. For improved performance of the AccessData Elasticsearch Windows Service, increase the heap size.

10. The *Configuration 2* dialog contains the following options:
    - **Discovery** - Selecting the default of *Multicast* allows the AccessData Elasticsearch Windows Service search to communicate across the network to other Elasticsearch services. If the network does not give permissions for the service to communicate this way, select *Unicast* and enter the IP address(es) of the server(s) that the AccessData Elasticsearch Windows Service is installed on in the *Unicast* host names field. Separate multiple addresses with commas.
    - **Node** - The Master node receives requests, and can pass requests to subsequent data nodes. Select both Master node and Data node if this is the primary system on which the AccessData Elasticsearch Windows Service is installed. Select only Data node if this is a secondary system on which the AccessData Elasticsearch Windows Service is installed. Click **Next**.
11. In the next dialog, click **Install**.
12. If the service installs properly, a command line window appears briefly, stating that the service has installed properly.
13. At the next dialog, click **Finish**.

Troubleshooting the AccessData Elasticsearch Windows Service

Once installed, the AccessData Elasticsearch Windows Service service should run without further assistance. If there are issues, go to **C:\Program Files\Elasticsearch\logs** to examine the logs for errors.